Review

Starting at the end: telomeres and telomerase in arthropods

Mauro Mandrioli¹✉, Valentina Monti¹,² and Gian Carlo Manicardi²

¹Dipartimento di Biologia, Università di Modena e Reggio Emilia, Modena, Italy
²Dipartimento di Scienze Agrarie e degli Alimenti, Università di Modena e Reggio Emilia, Reggio Emilia, Modena, Italy

*Corresponding author
e-mail: mauro.mandrioli@unimo.it

Abstract

Telomere composition and structure have been studied in several arthropods allowing us to better understand the evolution of such an important portion of the eukaryotic chromosomes. Genes coding for telomerase reverse transcriptase (TERT) have been sequenced and studied in few arthropod species only, where they resulted highly transcribed also in somatic tissues suggesting a different TERT regulation in species only, where they resulted highly transcribed also in somatic tissues suggesting a different TERT regulation in vertebrates. Contrary to the strict conservation of somatic tissues, telomere capping by reverse transcription based on telomerase, a highly conserved ribonucleoprotein present from unicellular eukaryotes to flowering plants and vertebrates (4). Nevertheless, not all the cells possess a transcriptionally active telomerase gene, and the progressive loss of the telomere repeats at the chromosome ends regulates both senescence and life span in somatic cells (4–6). The progressive loss of telomeric DNA in somatic cells can also act as a tumor suppressor mechanism making telomeres interesting also for understanding complex processes, such as aging and carcinogenesis, and explaining much of the current interest in the chromosomal ends (6, 7).

Telomere research became, in the last decades, a mainstream topic with papers facing telomere structure and functions from cell biology to oncology making impossible to review all the aspects in a comprehensive manner. In view of this assumption, we have decided to focus this review, as much as possible, on recent observations related to arthropods with canonical telomere/telomerase system.

Telomere composition and structure

Telomeres are generally composed of lengthy stretches of a simple repeat with a consensus sequence \((T\_A\_G\_T)\_x\). Telomeric DNA typically ends in a single-strand G-rich overhang of 50–300 nucleotides at the 3′ end that provides the basis for formation of non-Watson-Crick structures, such as G-quartets and t-loops (3) (Figure 1). In particular, t-loops protect telomeres by physically stitching the potentially vulnerable single-stranded G-strand terminus back into the double-stranded telomere sequence (8). According to literature data, t-loops also arrest the action of telomerase that extends telomeres, preventing their further lengthening (8).

The composition of telomeres may vary in eukaryotes, even if a strict conservation has been observed in some taxonomic groups so that the hexameric (TTAGGG), repeat is typical of vertebrates and other animals (9), the sequence (TTTAGGG), is common in plants (10), and the (TTAGG), telomeric repeat has been isolated in many of the main insect lineages and in other arthropods (11–16).

The length of the repeats has been evaluated in different species and varies not only between chromosomes (17), but also between species (11–13). In particular, telomere length evaluation, performed by telomere digestion with the exonuclease Bal31, indicated that the TTAGG terminal arrays of the lepidopterans Bombyx mori and Mamestra brassicae were about 6–9 kb long (11, 12). Further analyses, performed to evaluate the telomere length by digestion with restriction enzymes, revealed the presence of telomeric arrays longer than 21 kb in Pancrustacea (13). Similar results have been recently reported in the crustaceans Metapeneaus macleayi, Sagmariaus verreauxii, and Jasus edwardsii, where a telomere length of 10–20 kb has been assessed suggesting considerable lengths of the telomeric DNA in arthropods (18).

Introduction

Telomeres, from the Greek words for ‘end’ (telos) and ‘part’ (meros), are specialized structures constituting the end of each eukaryotic chromosome (1). They are involved in the protection of chromosome ends from erosion by exonucleases, and they avoid chromosome stickiness that could result in erroneous chromosome segregation during cell divisions (1, 2). Interestingly, as assessed in early work by McClintock and Muller and recently reviewed by Chan and Blackburn (3), normal chromosome ends lack the stickiness of chromosome breaks so that broken chromosomal ends often fuse with each other, whereas the telomeres do not (3).

As a consequence of the inability of DNA polymerase to fully replicate the 3′ end of the DNA strand (1, 2), telomeres are partially lost at each replication cycle in most of somatic cells (1). This loss can be faced by telomere elongation mediated by reverse transcription based on telomerase, a highly conserved ribonucleoprotein present from unicellular eukaryotes to flowering plants and vertebrates (4). Nevertheless,
Among insects, the (TTAGG)$_n$ repeat has been reported in Hymenoptera, Lepidoptera, Hemiptera, Trichoptera, and Megaloptera, but it seems to be absent in Ephemeroptera, Odonata, Dermaptera, and in the suborder Heteroptera (15). Moreover, the telomeric repeats are absent in the clade Antliophora (Diptera, Siphonaptera, and Mecoptera), where long repeated sequences (as in the non-biting midge Chironomus pallidivittatus) (19) or retrotransposable elements (as reported in the fruit fly Drosophila melanogaster) (20, 21) replace the canonical telomere-telomerase system, thus indicating that telomere elongation is telomerase-independent in some insects (21, 22). Furthermore, heterogeneity in the composition of telomeres has been observed in Coleoptera and Neuroptera (13), suggesting that the (TTAGG)$_n$ sequence is the ancestral motif of telomeres in insects. However, this telomeric repeat has been repeatedly lost or replaced with other motifs during insect evolution, also including alternative mechanisms of telomere elongation.

About 20 years ago, the (TTAGG)$_n$ sequence was identified as a component of telomeres and also in the crustaceans Gammarus pulex (14) and Penaeus semisulcatus (11), and more recently, Vitkova et al. (13) identified the (TTAGG)$_n$ repeat in several species belonging to Pancrustacea, Myriapoda, Chelicerata, and Pycnogonida suggesting that this motif represents an ancient telomeric sequence for arthropods.

Despite its ancient origin, the (TTAGG)$_n$ repeat seems to be derived from a more ancient sequence. In particular, as hypothesized by Vitkova et al. (13), the (TTAGGG)$_n$ motif seems to be much older than the (TTAGG)$_n$ sequence. Indeed, telomeres made by TTAGGG arrays are common in bilaterian animals (including Cephalochordata, Echinodermata, Onychophora, Platyhelminthes, Annelida, and Mollusca) so that it could be possible to hypothesize that the TTAGG motif evolved from the ancestral TTAGGG telomeric sequence.

**Synthesizing telomeres using the telomerase reverse transcriptase (TERT)**

Telomerase is a specialized reverse transcriptase consisting of a telomerase RNA-binding domain (TRBD), made up of α helices and two short β sheets, and the catalytic TERT domain capable of extending the 3′ end of chromosomes by adding telomeric repeats (23–25). Telomerase coding genes have been identified in the insects Apis mellifera, Acrystosiphon pisum, Tribolium castaneum, B. mori, and they encode for a 600- to 800-amino acid-long peptide with a 14–20% sequence identity to the vertebrate homologues (9, 26–28).

The four insect telomerases contain the same functional domains, but not all the motifs identified in the TERT of other eukaryotes. As first reported by Robertson and Gordon (17), the insect TERT presented seven conserved motifs (identified as 1, 2, A–E) defining the core RT domain, together with the TERT-specific T motif located immediately upstream to the core RT domain. The T motif is typical of TERT and absent in other reverse transcriptases not related to telomere synthesis (9). Different from vertebrate TERPs, insect telomerases miss the CP, GO, and QFP domains that have been identified in the N-terminal of the vertebrate TERT. These domains are also absent from Caenorhabditis elegans and Giardia lamblia telomerases (29) easily distinguishing vertebrate and invertebrate telomerase reverse transcriptases. No conserved domains specific to the insect TERTs have been identified (9).

TERT is highly regulated in human cells at both transcriptional and posttranscriptional levels so that most of the normal somatic cells lacks telomerase activity, whereas telomerase activation is observed in proliferating (such as activated lymphocytes) and cancer cells (30). Insect telomerases seem to be differently regulated as aphid TERT is highly expressed in different body parts, such as gut and head (26), in full agreement with Sasaki and Fujiwara (31) reporting telomerase activity in different organs and tissues of crickets and cockroaches. A somatic TERT expression was also evidenced in A. mellifera and B. mori where low amounts of telomerase mRNAs have been found in several tissues (27, 28). Interestingly, a weak telomerase activity was observed in different adult human tissues, but it is not sufficient to prevent telomere shortening. It could be therefore intriguing to further go in-depth in the study of TERT activity in insects, and in particular in A. mellifera and B. mori, in order to better comprehend the role of telomerase expression in the somatic tissues of these insects.

Consistent with the lack of a (TTAGG)$_n$ repeat, genes encoding for telomerase have been not identified in the dip- teran genomes (32, 33).
De novo synthesis of telomeres

Breakages of DNA double helices may result in chromosomal rearrangements, such as deletions, duplications, inversions, and translocations. However, in order to have recoverable chromosomal rearrangements, non-telomeric (broken) chromosome ends should not persist in the cell, as they induce cell cycle checkpoints arresting the cell cycle progression (34–37). Indeed, telomeres are involved in the chromosome stabilization, and broken non-telomeric chromosomal ends cannot replicate properly becoming highly unstable, and they have the propensity to fuse together (34–37).

Human telomeres are protected by the shelterin complex, which comprises six proteins that bind chromosomal ends in a sequence-dependent manner (36). Recent works showed that Drosophila telomeres are capped by a complex of fast-evolving proteins (called terminin) that is functionally analogous to shelterin (37). None of the terminin proteins is evolutionarily conserved outside the Drosophila species suggesting that flies rapidly evolved terminin to bind chromosome ends in a sequence-independent fashion probably slightly before the loss of telomerase (37).

Telomere stabilization may also involve the addition of repetitive telomeric sequences at the breakpoints by telomerase (de novo telomere synthesis). Hence, the addition of telomeric repeats results in the stabilization of the new chromosome end, and it allows the resumption of cell cycling (34, 35). Stabilization of broken chromosome ends by telomere sequence addition has been observed in many organisms, from yeast to man (34, 35), but until now in three insect species only (the dipteran D. melanogaster and hemipterans Planococcus lilacinus and A. pisum) (21, 26, 38).

The presence of de novo synthesis is particularly interesting in aphids (Figure 2) and coccids as they both have holocentric/holokinetic chromosomes possessing centromeric activity spread along the whole chromosomal axis (39, 40). This peculiar chromosome feature, coupled with the de novo telomere synthesis stabilizing the breakpoints, allows a proper stabilization of chromosomal fragments assuring their inheritance during cell divisions.

Looking below telomeres: the subtelomeric regions

Different from the conservation of the telomeric sequences, insect subtelomeric regions are more polymorphic and variable in composition. As a general rule, repetitive telomere-associated sequences (TAS) have been commonly found in the subtelomeric region of various insect species, such as the 169-bp MpR satellite DNA sequence in the aphid  

Figure 2  The aphid M. persicae is one example of an insect telomere with canonical telomerase and (TTAGG) repeat. FISH with the FITC-labeled (TTAGG) probe (A, B) evidenced that each chromosomal end consists of an array of the TTAGG motif, not only in a standard karyotype (A), but also in metaphase plates where a fragmentation occurred suggesting that a de novo synthesis of telomeres occurred (B). In view of the presence of the MpR subtelomeric satellite at each autosomal subtelomeric end (C), it has been possible to distinguish standard (D) and neo-synthesized telomere (E) by fiber FISH. In standard telomere, FISH on DNA fibers stained with DAPI (in blue) showed the presence of the TTAGG array (in red due to the use of a TRITC-labeled telomeric probe) near the cluster of the MpR subtelomeric satellite (labeled in green in view of the use of a FITC-labeled probe) (D). The MpR subtelomeric array is absent from de novo telomeres (E). Asterisks indicate the chromosomal ends involved in fragmentation. Arrows indicate X chromosomes.
Myzus persicae (41) (Figure 2) and the highly conserved 9-kb long terminal unit (LTU) identified in the Taiwan cricket Teleogryllus taiwanemma (42). Both these repetitive sequences were located at almost all the subtelomeric regions and were species-specific or, at most, present in a few highly related species. Indeed, the 169-bp MpR subtelomeric satellite has been found in M. persicae, Myzus antirhini, and Myzus curvus, but absent in other aphid species (41). Similarly, T. taiwanemma LTUs resulted absent in other crickets, including the Japanese field cricket Teleogryllus emma that is thought to be one of the species closest to the Taiwan cricket T. taiwanemma (42). As a whole, it emerges that TAS sequences have been rapidly amplified in subtelomeric regions by recent evolutionary events and may act as a backup system to prevent telomere shortening when the telomerase activity is blocked (43).

Interestingly, the TASs identified up until now bear a structural resemblance to Chironomus TA repeats (44), which evolved from telomeric repeat sequences and truncated retrotransposons (19, 44, 45) suggesting that retrotransposons could be common elements located below telomeres and their evolution shaped the structure of the subtelomeric regions.

The presence of non-LTR retrotransposons has been frequently reported in insects and TRAS and SART retrotransposons have been isolated from the subtelomeric regions of the lepidopterans B. mori, Dictyoploca japonica, Samia cynthia ricini, and M. brassicae (11, 12, 46). Furthermore, TRAS elements have been annotated in the genome projects of the aphid A. pisum (47) and the beetle T. castaneum (48).

More than 2000 copies of non-LTR retrotransposons belonging to the TRAS and SART families have been identified in B. mori proximally to the (TTAGG)ₙ repeats. TRAS and SART were abundantly transcribed and actively retrotransposed into the TTAGG telomeric repeats in a highly sequence-specific manner (11). Surprisingly, no insertions of non-LTR or any other retrotransposons have been reported in the subtelomeric regions of A. mellifera (9, 27).

Subtelomeric regions are therefore composed of complex patchwork of different moderately and highly repeated sequences, interspersed into degenerate telomeric repeats (49). Moreover, the subtelomeric regions of most organisms are dynamic with frequent turnover and exchange of sequences (49).

Despite their sequence variation, arthropod chromosome ends are similar in structure suggesting the existence of shared functional constraints that require this chromosomal region (49). At present, the functional roles of subtelomeric regions have been not deeply studied in insects, but Drosophila, in spite of its exception in the telomere structure, is furnishing new insights about insect telomeres and subtelomeres (50). For instance, fly TAS sequences are involved in a silencing phenomenon (called telomeric position effect) that is due to a specific chromatin conformation of the TAS located in the subtelomeric regions of chromosomes (51).

TAS elements can also regulate telomere length in different ways (52). Hence, telomere growth is likely to be regulated by the organization of the subtelomeric chromatin so that at each telomere, the telomeric complex and subtelomeric chromatin cooperate to form a unique higher-order chromatin structure that controls telomere length (52). Last, TAS elements can act as transcription initiation sites for telomere repeat-associated transcripts that can negatively regulate the telomerase-dependent telomere elongation (53).

To be continued... over the end

In the last years, several studies faced different aspects of telomere structure and genetics, including a large number of papers that analyzed chromosomal ends in non-model organisms. This approach leads to a much deeper understanding of the origin, nature, and evolution of telomeres and their maintenance systems. Recently, telomeric repeat-associated siRNAs (tel-siRNAs) have been isolated in plants, and they resulted conserved in a wide range of crop species showing that tel-siRNAs have a potential regulatory role in telomere dynamics (54). The presence of tel-siRNAs associated to telomeric chromatin has not been deeply analyzed in insects, with the exception of D. melanogaster (53), making the comprehension of non-coding RNA involvement in the regulation of telomere functioning a new frontier in the telomere biology. Despite several decades of studies, new discoveries about telomere epigenetics are clearly showing that telomere studies are quite far from the end.

Acknowledgements

This work was supported by a grant ‘Experimental approach to the study of evolution’ from the Department of Animal Biology of the University of Modena and Reggio Emilia (M.M.).

References


Received March 9, 2012; accepted June 15, 2012

Mauro Mandrioli obtained his degree in Biological Sciences in 1996 and four years later his PhD in Animal Biology at the University of Modena. Since 2010 he is associate professor in genetics at the University of Modena and Reggio Emilia, after working from 2001 to 2010 as researcher in genetics. His main interests are related to the cytogenetics and epigenetics of pest crop insects with holocentric chromosomes.

Gian Carlo Manicardi obtained his degree in Biological Sciences in 1984 and his PhD in Animal Biology at the University of Modena in 1989. Since 2002 he is associate professor in genetics at the University of Modena and Reggio Emilia, after working from 1990 to 2002 as researcher in genetics. His main interests are related to genome organization in pest crop insects chromatin anomalies and to DNA damages in mammalian spermatozoa in relation to the use of pesticides and/or other endocrine disrupters.

Valentina Monti obtained her degree in Biology in 2010 at the University of Modena and Reggio Emilia and she is currently PhD student at the University of Modena and Reggio Emilia, where she is working on chromosomal rearrangements in aphids, with particular regard for the peach potato aphid *Myzus persicae*.